(12) UK Patent Application (19) GB (11) 2 365 296 (13) A

(43) Date of A Publication 13.02.2002

(21) Application No 0018491.1

(22) Date of Filing 27.07.2000

(71) Applicant(s)
Cambridge Consultants Limited
(Incorporated in the United Kingdom)
Science Park, Milton Road, CAMBRIDGE, CB4 4DW,
United Kingdom

(72) Inventor(s)
Roger Fane Sewell
Mark St John Owen
Stephen John Barlow
Simon Paul Long

(74) Agent and/or Address for Service
Beresford & Co
2-5 Warwick Court, High Holborn, LONDON,
WC1R 5DH, United Kingdom

(51) INT CL[7]
H04H 1/00 , G11B 20/00 , H04N 5/913

(52) UK CL (Edition T)
H4R RPX R22B R22V
G5R RHB
H4P PEUM

(56) Documents Cited

GB 2348028 A	GB 2343818 A
EP 0905967 A	EP 0891071 A
EP 0828372 A	WO 96/42151 A

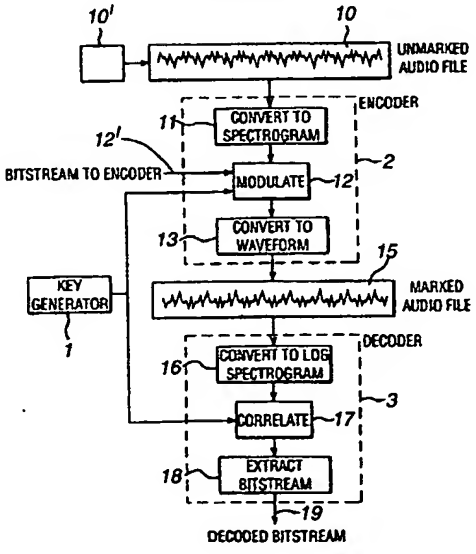
(58) Field of Search
UK CL (Edition S) G5R RHB , H4F FBB , H4P PDCFX PDX PEUM , H4R RPTS RPX
INT CL[7] G11B 20/00 , H04H 1/00 , H04N 1/32 5/913 7/08
online: EPODOC,WPI,JAPIO

(54) Abstract Title

Encoder/decoder for watermarking a covertext signal

(57) To watermark a covertext comprising audio or video signals 10, a two-dimensional pattern of predetermined size is used to generate at least one key 1 which is used by an encoder 2 to watermark the covertext and which can be used subsequently by a decoder 3 to decode the resulting stegotext 15. In the encoder, multiples of the key(s) are added or subtracted to blocks of the covertext signal transformed into a power spectrogram 11 by fast Fourier transformation and rectangular polar conversion of the covertext signal to represent a desired code. The stegotext is then obtained by polar rectangular conversion and inverse fast Fourier transformation. In the decoder 3, the stegotext is transformed into the log power spectrogram domain by fast Fourier transformation and rectangular polar conversion. Two-dimensional correlation 17 is then performed between the original key(s) and blocks of the transformed stegotext so as to generate a correlation function from which data bits representing the code are extracted at 18. The origin of the covertext can thus be identified in an imperceptible way, and unauthorised reproduction of the covertext may be prevented.

FIG. 2
SYSTEM OVERVIEW



GB 2 365 296 A

FIG. 1

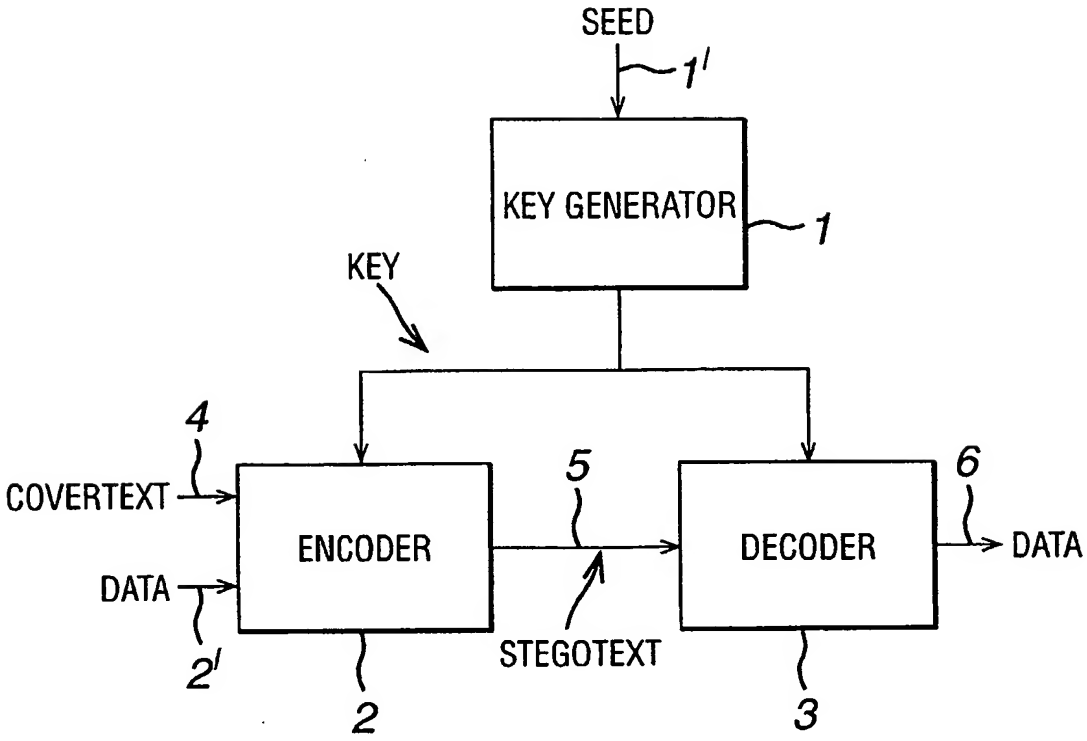

SEED
1'
KEY GENERATOR
1
KEY
4
COVERTEXT
DATA
2'
ENCODER
2
5
STEGOTEXT
DECODER
3
6
DATA

FIG. 7

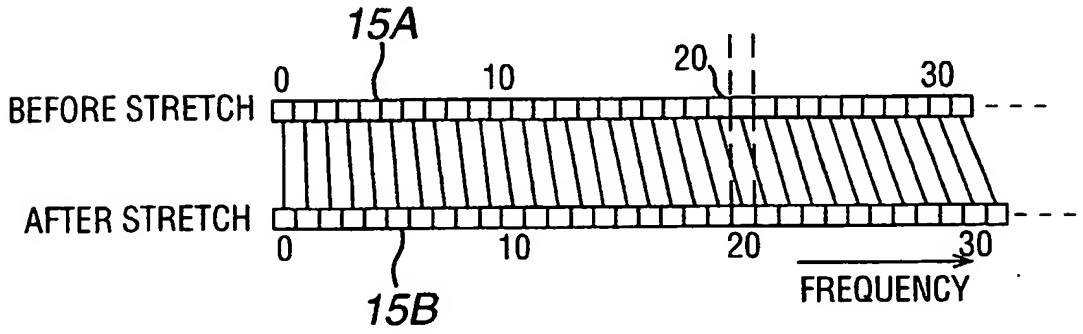

15A
0
10
20
30
BEFORE STRETCH
AFTER STRETCH
0
10
20
30
FREQUENCY
15B

FIG. 2
SYSTEM OVERVIEW

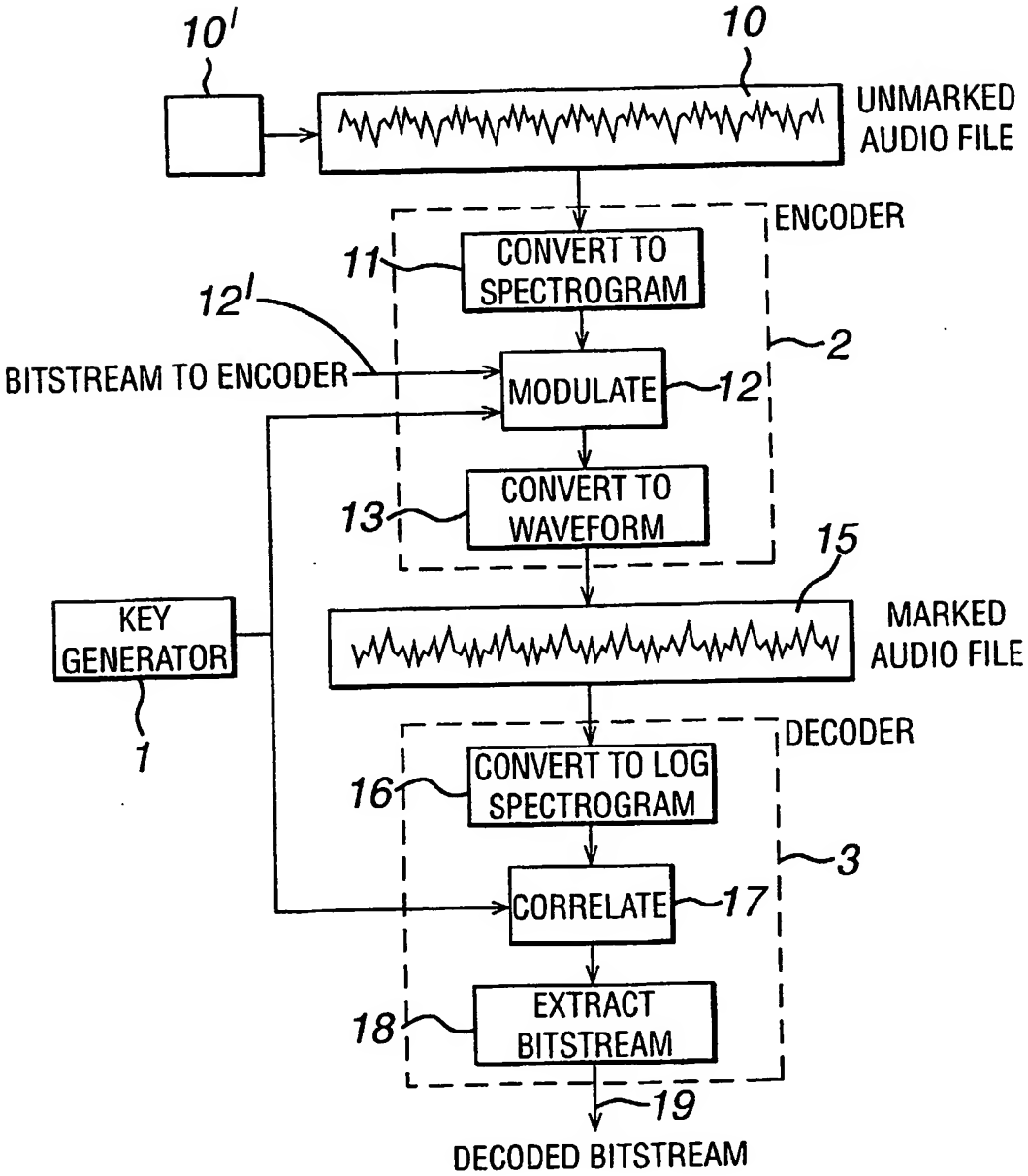


FIG. 3

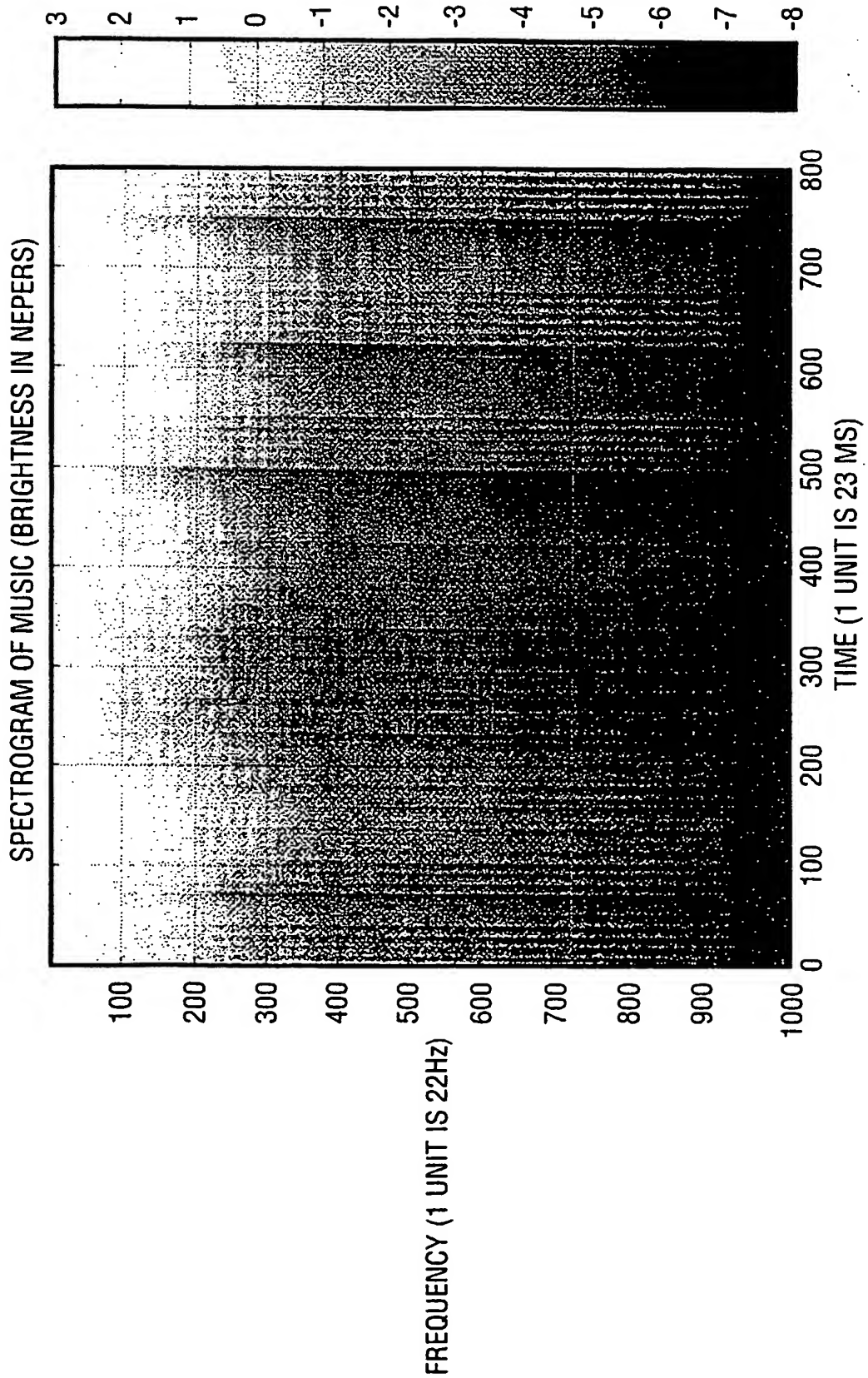

SPECTROGRAM OF MUSIC (BRIGHTNESS IN NEPERS)
3
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
FREQUENCY (1 UNIT IS 22Hz)
100
200
300
400
500
600
700
800
900
1000
0
100
200
300
400
500
600
700
800
TIME (1 UNIT IS 23 MS)

FIG. 4 OVERLAPPING MODULATION PATTERNS

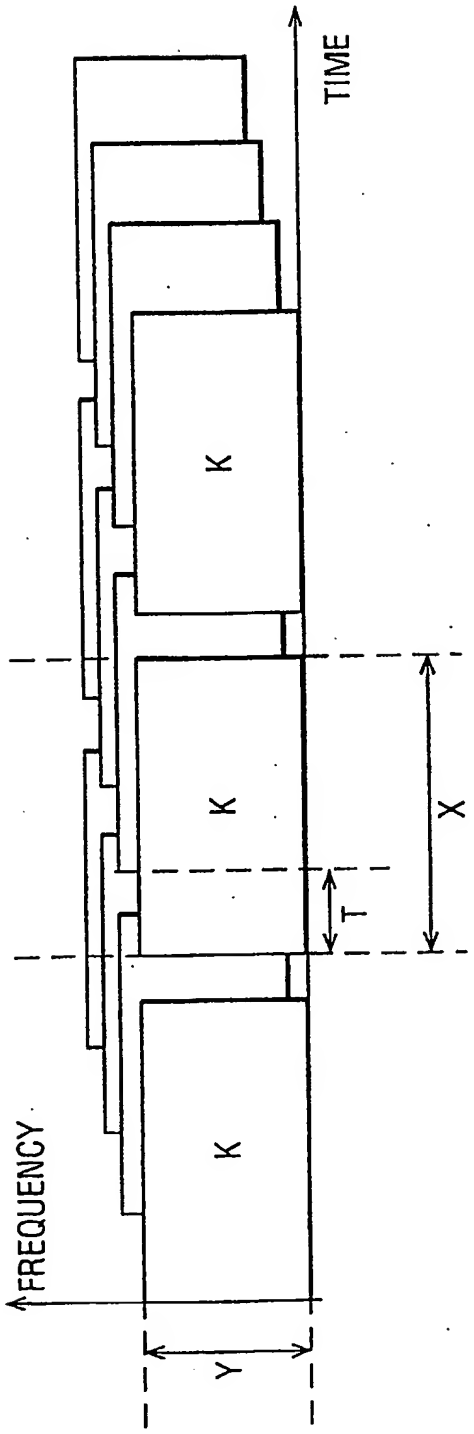


FIG. 5
CONVOLUTIONAL ENCODER

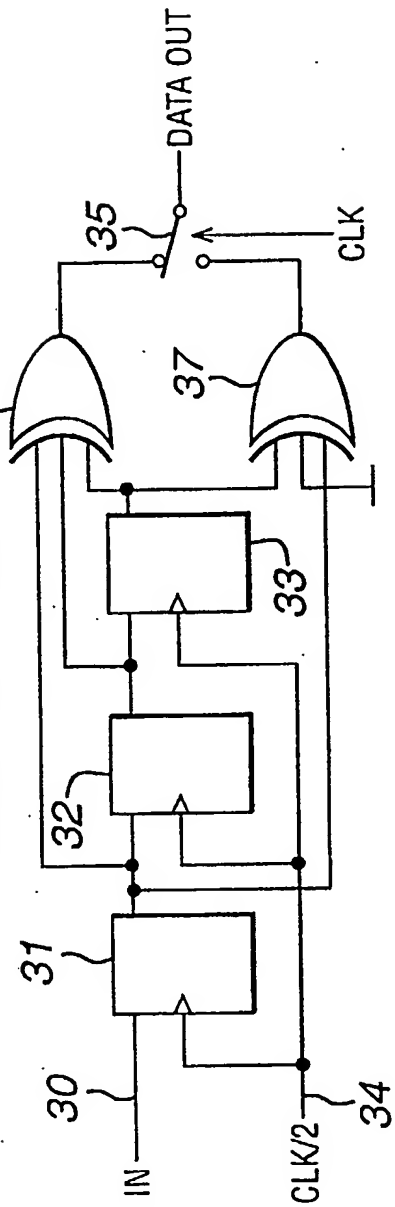


FIG. 6

MODIFICATION OF POWER SPECTROGRAM USING WINDOWED FFTs

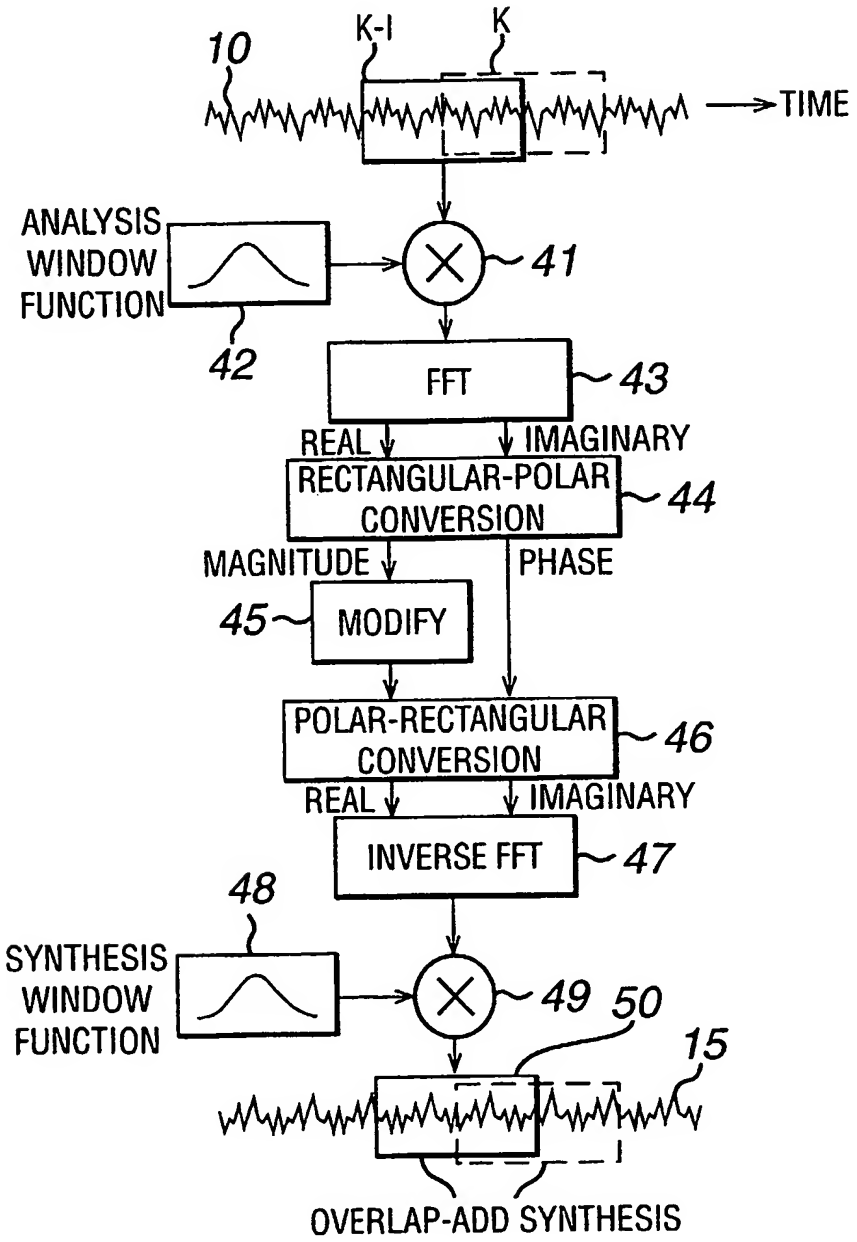


FIG. 8

KEY GENERATION FILTER PARAMETERS

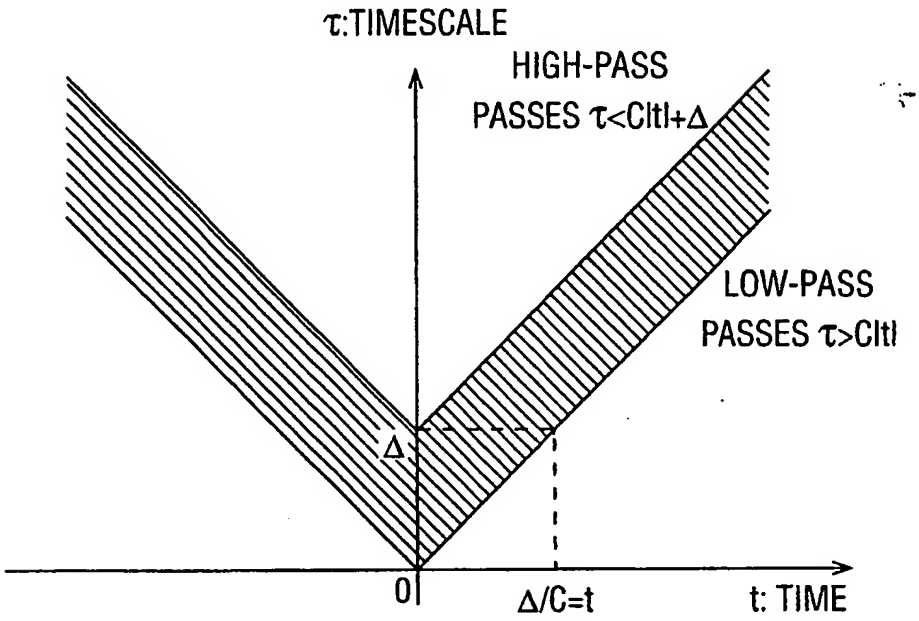


FIG. 9

FILTER CHARACTERISTICS OF CONSECUTIVE KEYS

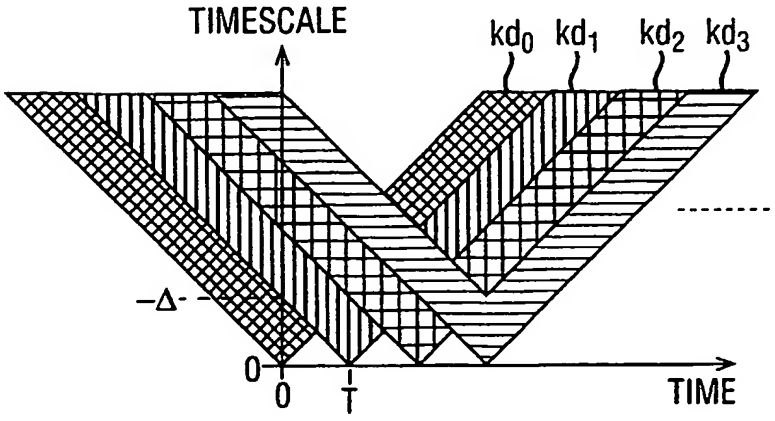


FIG. 10

ONE-DIMENSIONAL WHITE NOISE SIGNAL
WITH SWEPT BAND-PASS FILTER

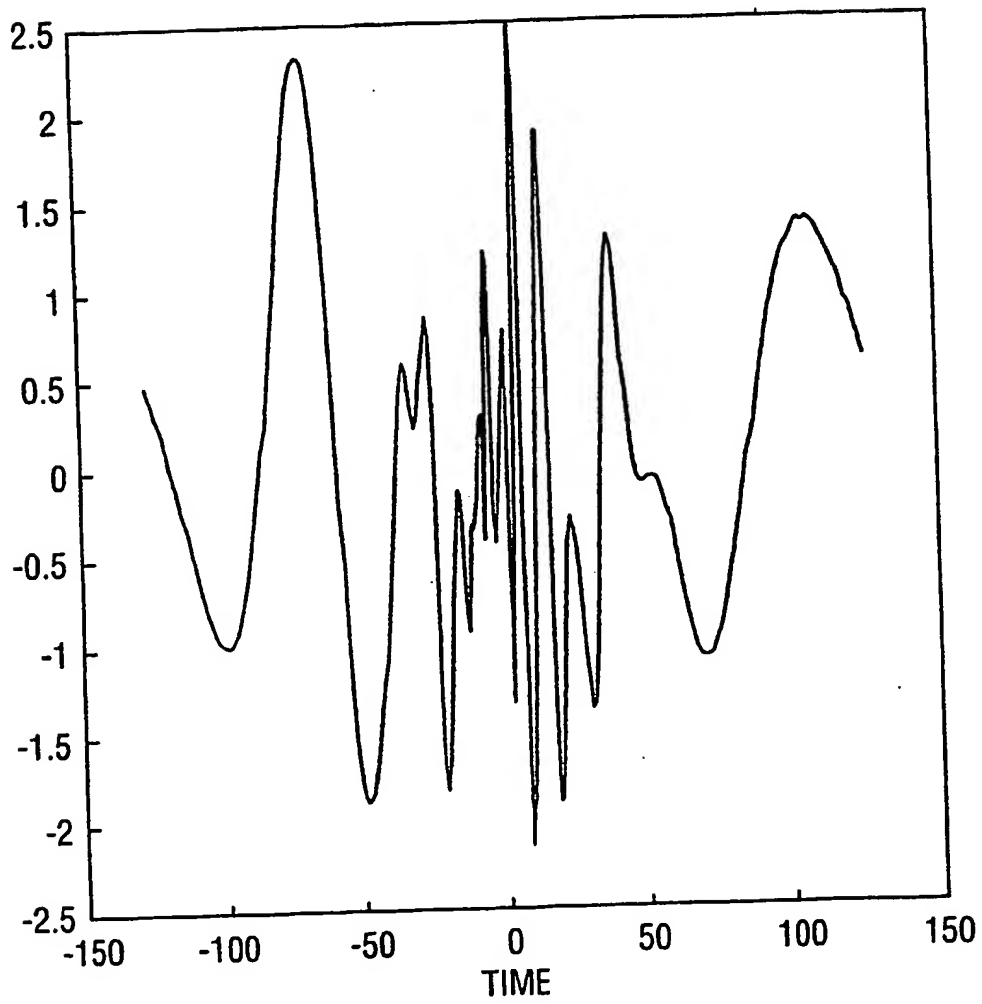


FIG. 11

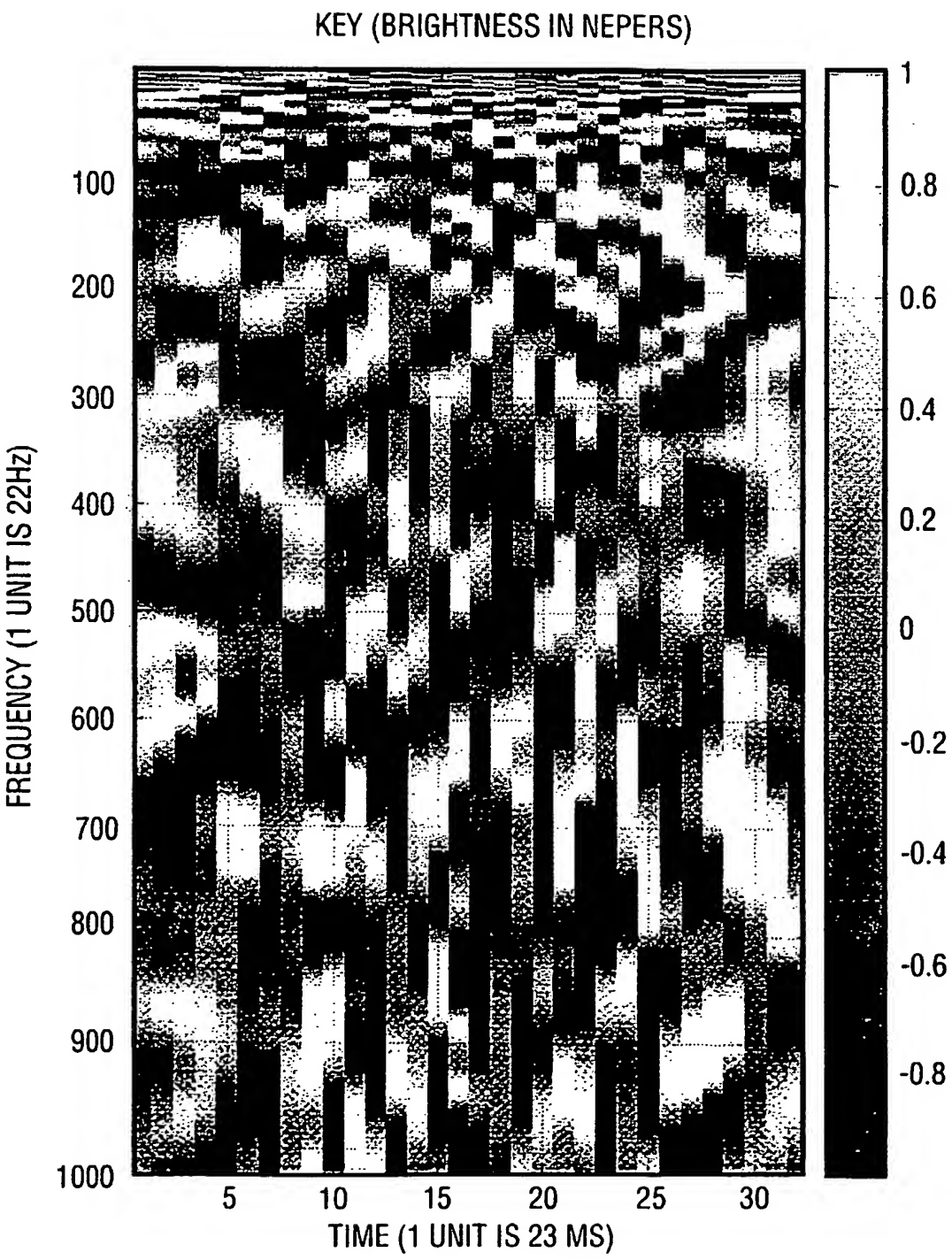

KEY (BRIGHTNESS IN NEPERS)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
FREQUENCY (1 UNIT IS 22Hz)
100
200
300
400
500
600
700
800
900
1000
5
10
15
20
25
30
TIME (1 UNIT IS 23 MS)

FIG. 12 EFFECT OF TIME AND FREQUENCY STRETCH ON AMPLITUDE OF CORRELATION PEAK

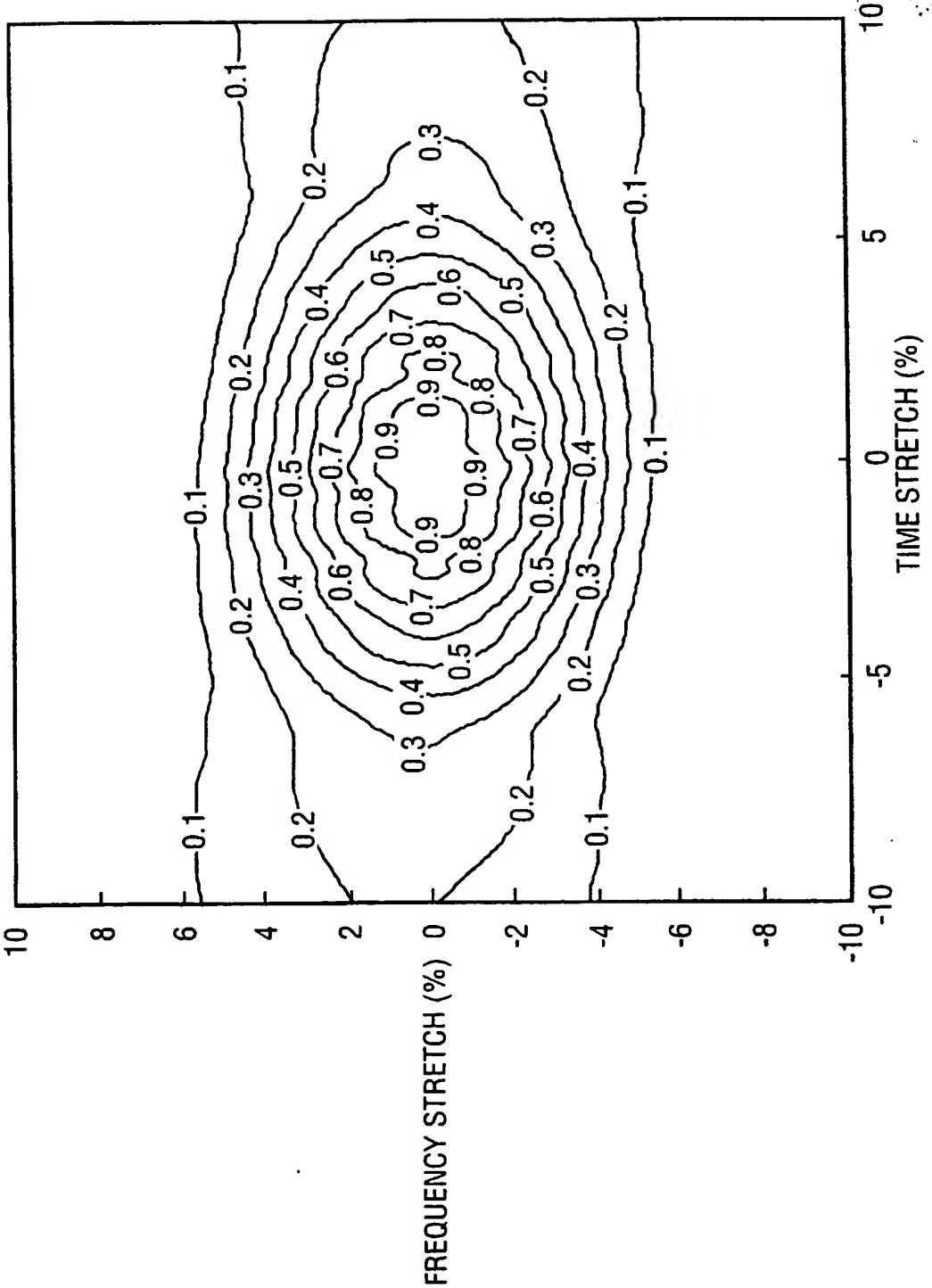


FIG. 13 ENCODER BLOCK DIAGRAM

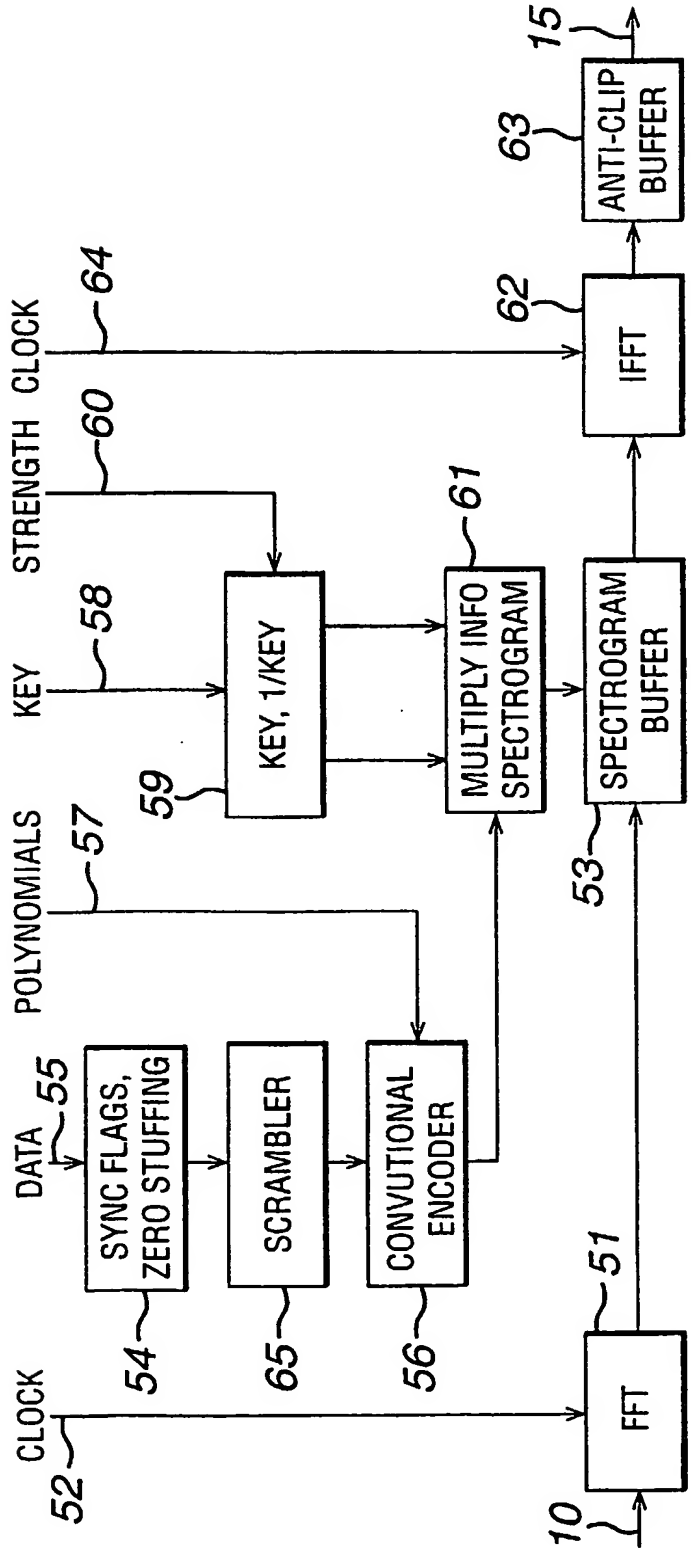


FIG. 16

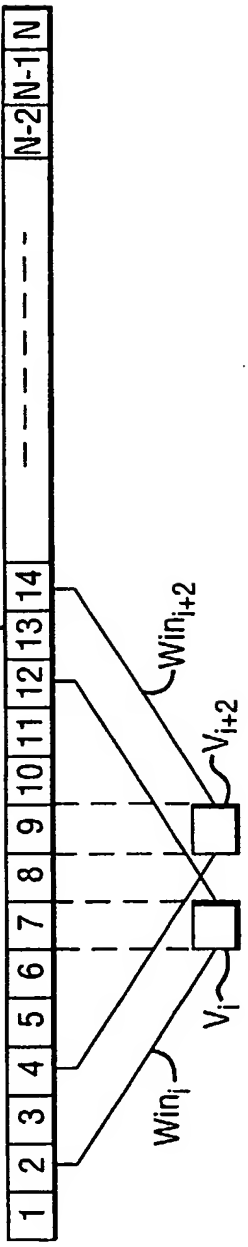


FIG. 14 DECODER BLOCK DIAGRAM

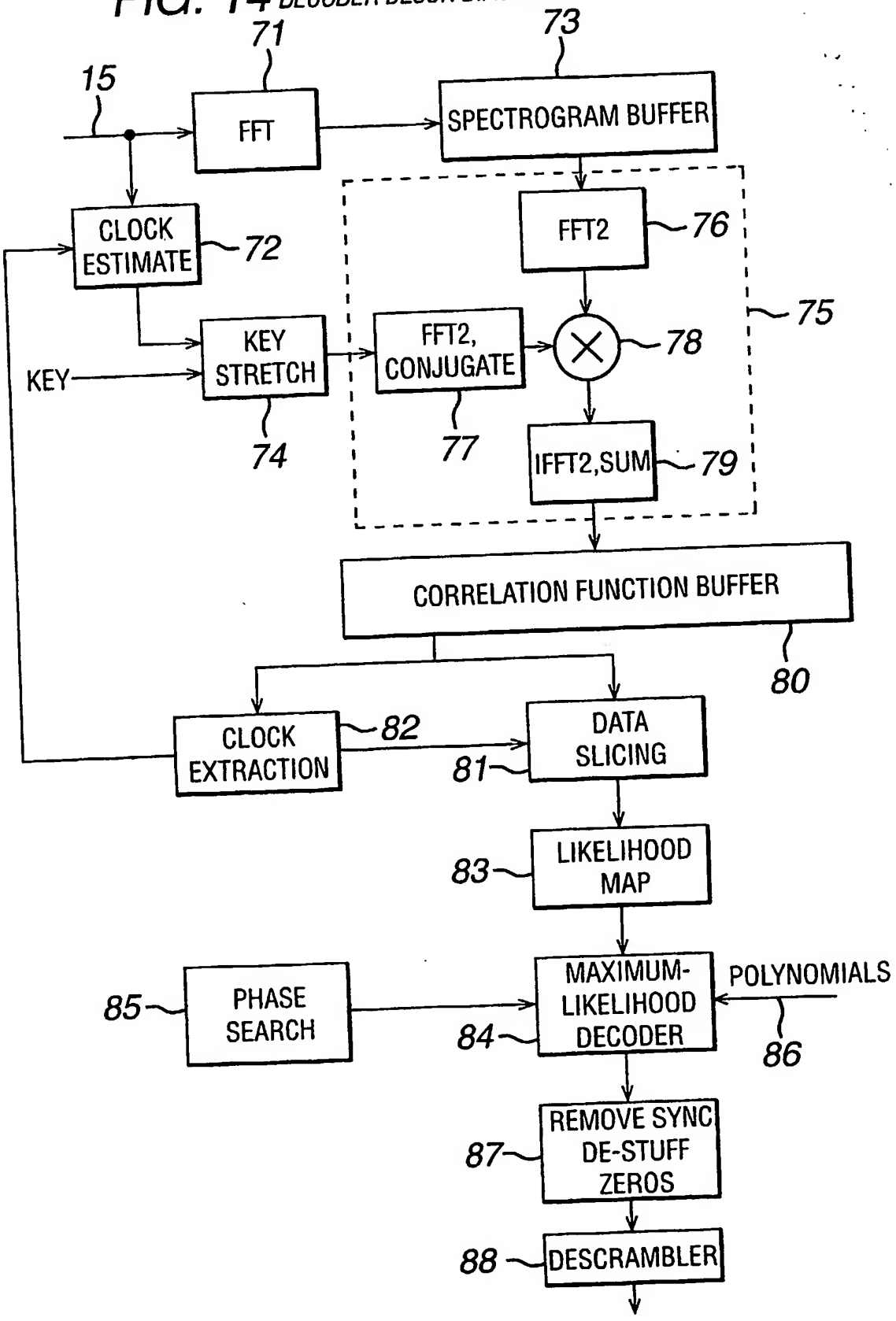
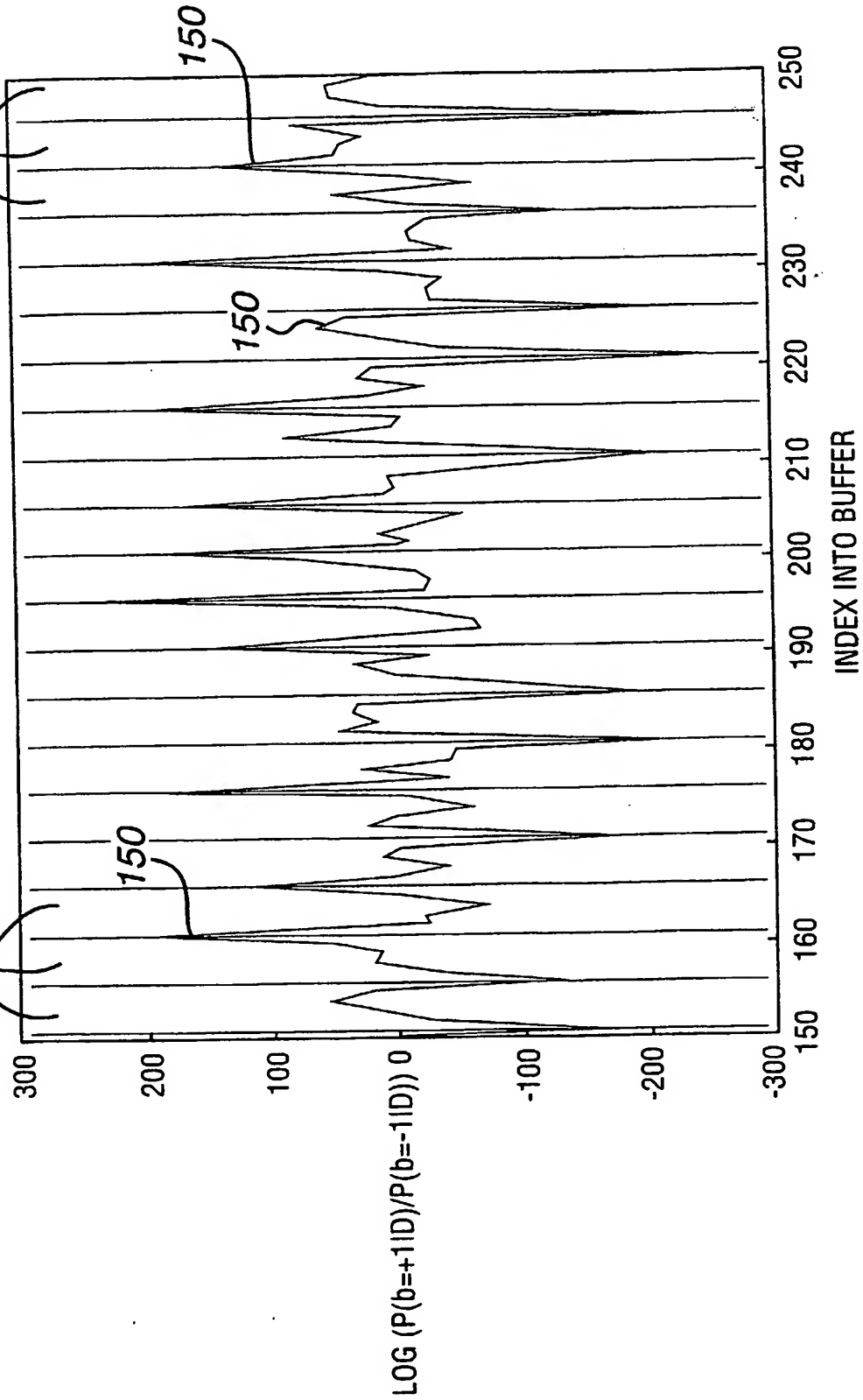


FIG. 15


150
151
150
150
151
300
200
100
0
-100
-200
-300
150
160
170
180
190
200
210
220
230
240
250
LOG (P(b=+1|D)/P(b=-1|D))
INDEX INTO BUFFER

POWER SPECTROGRAM ENCODER

The present invention concerns watermarking analog or digital signals. It will be appreciated that whilst the signals may be video or data signals the present invention is particularly, though not exclusively, concerned with watermarking analog or digital audio signals.

The term "watermarking" is intended to cover the procedure of adding data to a main signal so that the added data does not affect the main purpose of the main signal. The main signal is often referred to a "covertext" and the signal with the added watermarking data is often referred to as a "stegotext". Thus in the case of an audio signal the presence of the added data in the stegotext is intended to be virtually imperceptible to a listener when the stegotext is reproduced. However the presence of the added data in the stegotext enables, if the user has the appropriate decoding equipment, to identify the origin of the covertext. If the user's equipment is provided with suitable circuitry he or she may be prevented from reproducing the main data carried

in the original covertext signal if the watermark data recovered does not match the equipment. Additionally a user has to be able to reproduce a covertext.

Such techniques obviously have great potential with regard to musical recordings. As a result a substantial amount of effort has been put into the problem of watermarking audio signals in such a manner that a person who is entitled to listen to the stegotext does not have his or her enjoyment impaired by spurious sounds caused by the added coded data.

Alternatively it is important that the watermarking should be sufficiently robust both to remain effective after the various types of conventional signal processing to which recorded and transmitted audio material can be subjected and also to be able to resist direct attempts to eliminate or render ineffective the added coded data.

An apparatus and method for watermarking analog signals is disclosed in International Patent Specification No. WO98/53565, and this specification discloses a number of the techniques which have been employed to watermark signals.

One method of watermarking proposed in this prior

published specification involves measuring the short-term autocorrelation function of the audio signal and then adding an additional signal which is hard to hear and which changes the value of the short-term autocorrelation function at some specific delay or delays to produce a specific waveform which carries data at a low rate. The actual modulation of the data onto this waveform can be done by using any of a number of suitable modulation techniques. At the reception end of the apparatus a watermark reader (or decoder) measures the short-term auto correlation function of the stegotext and applies a demodulation appropriate to the modulation technique used. Provided that the reader can utilise the data which was initially used to modulate the autocorrelation function the added coded data can be removed from the stegotext.

However the short-term autocorrelation function of many audio signals can be easily altered to be arbitrarily close to zero at arbitrarily long delays without altering the sound of the basic audio. It is accordingly possible to attack the watermarked signal in a relatively simple manner so as to nullify the effect of

the watermarking.

The present invention is concerned with providing a watermarking system which is not subject to the above defect and also to provide a decoder for decoding watermarked signals.

In accordance with one aspect of the present invention there is provided an encoder for encoding a covertext signal to generate a stegotext, the encoder comprising first transformation means for carrying out a Fast Fourier Transform and rectangular polar conversion of the covertext signal so as to transform the covertext signal into a log power spectrogram; means for providing at least one key, the or each key being in the form of a two-dimensional pattern of predetermined size; a multiplier for adding or subtracting in the log power spectrogram domain multiples of the key, or multiples of one or more of the keys if there is a plurality of keys, to blocks of the transformed covertext signal; means for controlling the addition or subtraction of the key or keys by the multiplier in accordance with data representing a desired code, and second transformation means for carrying out polar rectangular conversion and

inverse Fast Fourier Transformation of the modulated covertext signal to generate a stegotext.

In accordance with a second aspect of the present invention there is provided a method of encoding a covertext signal to generate a stegotext, the method comprising carrying out a Fast Fourier Transform and rectangular polar conversion of the covertext signal so as to transform the covertext signal into the power spectrogram domain; providing at least one key, the or each key being in the form of a 2-dimensional pattern of predetermined size; adding or subtracting in the log power spectrogram domain multiples of the key, or multiples of one or more of the keys if there is a plurality of keys, to segments of the transformed covertext signal; controlling the addition or subtraction of multiples of the key or keys at the addition/multiplication step in accordance with data representing a desired code; and carrying out polar rectangular conversion and inverse Fast Fourier Transform of the modulated covertext signal to generate a stegotext.

In accordance with a third aspect the present

invention provides a decoder for decoding a stegotext signal watermarked by the method set out hereinbefore to obtain the watermarking code, the decoder comprising transformation means for carrying out Fast Fourier Transformation and rectangular polar conversion of the stegotext signal so as to transform the stegotext signal into the log power spectrogram domain; means for providing the key or keys with which the original covertext signal was encoded; means for carrying out 2-dimensional correlation between the key or keys and blocks of the transformed stegotext signal which are of the same length as the key so as to generate a correlation function representing the correlation between the key and the stegotext, and means for extracting the data bits representing the code from the correlation function.

In accordance with a fourth aspect the present invention provides a method of decoding a stegotext signal watermarked to obtain the watermarking code, the method comprising carrying out Fast Fourier Transformation and rectangular polar conversion of the stegotext signal so as to transform the stegotext signal

into the log power spectrogram domain; providing the key or keys with which the original covertext signal was encoded; utilising the key or keys to obtain a 2-dimensional correlation function between the key or keys and sequential blocks of the transformed stegotext signal which are of the same length as the key or keys so as to generate a correlation function representing correlation between the key and the stegotext, and extracting the data bits representing the code from the correlation function.

In order that the present invention may be more readily understood embodiments thereof will now be described by way of example and with reference to the accompanying drawings, in which:

Figure 1 is a block diagram of a system for encoding and decoding a covertext signal with additional data so as to generate a stegotext;

Figure 2 is a block diagram of an encoder and decoder which can be used in the embodiment of Figure 1 to generate and decode a stegotext;

Figure 3 is an illustration of the power spectrum of a segment of music;

Figure 4 is a diagram illustrating the overlapping of modulation patterns when the power spectrogram is modified;

Figure 5 is a block diagram of a convolution encoder;

Figure 6 is a block diagram of an encoder/decoder shown in greater detail than the embodiment of Figure 2;

Figure 7 is a diagram illustrating a time-stretch attack on a stegotext;

Figure 8 is a graph illustrating filter parameters used in accordance with an embodiment of the present invention in the system of Figure 1;

Figure 9 is a graph showing the filter characteristics of consecutive keys;

Figure 10 is a graph showing the one dimensional white noise signal with a swept band-pass filter;

Figure 11 is a graph illustrating the result of two dimensional white noise with swept filters in each direction;

Figure 12 is a graph illustrating the effects of stretch on correlation;

Figure 13 is a block diagram of one embodiment of an

encoder in accordance with the present invention;

Figure 14 is a block diagram of an embodiment of a decoder according to the present invention.

Figure 15 is a graph indicating the contents of a buffer of the decoder of Figure 14; and

Figure 16 is a diagram illustrating the operation of a maximum-likelihood convolutional-code decoder which forms part of the decoder of Figures 14.

Referring now to Figure 1 the basic system consists of a key generator (1), an encoder (2) and a decoder (3). The key generator (1) produces a pseudo-random key based on an integer seed value input at (1'). The encoder (2) marks a music file input at (4) as a cover text with data using the key to generate a stegotext. The data is input into the encoder (2) at (2'). The decoder (3), receiving the stegotext, over a transmission line (5), reads back the data from a marked file again using the key and outputs the recovered data at (6). The same key must be used in the encode and decode operations to ensure that the data are read back correctly. The key can, of course, be regenerated when needed from the seed, so the seed value is all that is required to decode a marked

file. The transmission line (5) can, of course, take a wide variety of forms. Thus the stegotext could be recorded on any suitable medium or transmitted by radio, fibre cables or the like. Hereinafter any unmarked file will be referred to as the covertext and a watermarked file as the stegotext. Whilst the present embodiment is described in relation to its use with music it will be appreciated that the techniques and apparatus described can be used in non-musical situations such as speech or video data.

Figure 2 of the accompanying drawings shows a block diagram of a more detailed embodiment in accordance with the present invention. In this Figure the covertext is an unmarked audio file which is shown at (10) The source of the audio file is indicated at 10'. This can be a microphone picking up a live event, a recording such as a tape or disc or a signal which has been transmitted by radio or the Internet. This audio file is input to the encoder (2) and in circuit (11) is converted into a power spectrogram. The reason for this conversion is as follows. It is not feasible to convey information in the phase components of the stegotext. The human ear is

essentially insensitive to phase, a fact exploited by some compression algorithms. Accordingly a watermarking technique that depends on phase is unlikely to be robust to compression. Moreover it is possible to process an audio file, scrambling the phases of its frequency components, by applying a random "group delay" to the file. Such processing, which is not computationally intensive, will in general destroy any particular wave shape present in the audio file. Thus watermarks which depends on wave shape, that is on the time domain form of the signal, can be rendered unreadable by this processing.

Accordingly in the present invention it is proposed to carry out the watermarking of a covertext by using the power spectrum of the covertext. Thus only the magnitude of each frequency component in the covertext is modified and the phase of each frequency component is preserved throughout the marking process. Phase information is discarded in the decoder. This procedure will now be described in greater detail.

In order to calculate the power spectrogram of the covertext the covertext is divided into blocks 2Y samples

long that overlap by half their length. Thus a new block starts every Y samples. In the present embodiment, which as described is designed for audio files with a sample rate f_s = 44100 Hz, Y is set to 1024.

Each block is multiplied by a window function, known as the analysis window, and the Fourier transform of the windowed block is calculated. The purpose of the window function is to ensure that the sample values taper off towards zero at either end of the block, avoiding a discontinuity. The Fourier transform treats the block as the repeating unit of a periodic function. Since the windowed block consists of real samples, its Fourier transform is conjugate symmetric with respect to positive and negative frequencies. The negative frequency components carry no additional information and can therefore be discarded.

Each Fourier coefficient is a complex number whose magnitude represents the amplitude of the corresponding frequency component and whose argument represents its phase. When the phase information is discarded, what remains is the power spectrum of the signal. In a strict sense the power spectrum is obtained by squaring the

magnitude of each Fourier coefficient.

When a number of consecutive power spectra are placed alongside one another, a grid of values is formed: one axis, conventionally vertical, represents frequency whilst the other, conventionally horizontal, represents time. This grid is the power spectrogram of the audio sample. Figure 3 of the drawings is an example of a power spectrogram taken from a segment of music. In this figure the values in the grid are shown as various shades of grey. The right hand column running from -8 to 3 is a scale against which the brightness levels of the spectrogram can be matched so that the spectrogram can be evaluated.

The choice of Y determines the resolution of the spectrogram. In the frequency direction, the resolution is $f_s/2Y$; in the time direction, the resolution is Y/f_s. In the present embodiment these values are 21.5 Hertz and 23.2ms respectively. The axes of Figure 3 are measured in these units.

Whilst it may appear to be difficult satisfactorily to reconstruct an audio waveform from its power spectrogram it is possible if phase information is

retained. The spectrogram data can be returned into the time domain with an inverse Fourier transform, overlapped in the same manner before, and added together.

In order to watermark the covertext from which the spectrogram was obtained it has been discovered that as long as modifications to the spectrogram are small and as long as the original phase information is retained the above described method recreates a satisfactory audio waveform. It is to be observed that the reconstructed time-domain segments are no longer guaranteed to taper to zero at either end; the subjective quality of the final waveform is therefore improved if the segments are windowed with the synthesis windows as described before being added together. The analysis and synthesis windows must be chosen to ensure that there is no overall amplitude modulation through the system. In the present embodiment each of these windows is the square root of a raised-cosine function.

In Figure 2 the modulation of the spectrogram is carried out in a circuit generally indicated at (12) in response to the bit stream to be encoded.

Finally in block (2) circuit (13) returns the

modulated power spectrograms to the time domain and synthesises these so as to convert them into the stegotext. In Figure 2 the stegotext is indicated at 15.

The decoder (3) comprises a circuit (16) for converting the stegotext to a log spectrogram, a circuit (17) utilising the key to correlate the log spectrogram so as in circuit (18) to extract the bit stream representing the watermark code and which is output at (19).

It has been discovered that the extent to which an element of a power spectrogram of an audio signal can be modulated without audible effect is roughly proportional to its original level. Thus in decibel terms additions or subtractions may be made to the power spectrogram up to a fixed amount. The amount of modulation that is perceptible depends on the listening environment, but is typically around 1dB. Accordingly in the present embodiment the watermarking process is carried out in the "log power spectrogram domain" and consists in making additions or subtractions to the power spectrogram in accordance with the key generated by key generator 1 and the data to be encoded as the watermark. The data is

input at 12'.

Since a greater degree of modulation can be applied to spectrogram elements with larger magnitude the information carried in those elements will be less susceptible to noise than in elements with smaller amplitude. However it is impossible to know beforehand which these elements will be. Thus the watermarking scheme being described is prepared to exploit whichever elements the covertext makes available for carrying information. Thus in the present embodiment each spectrogram element in circuit (12) is modulated so as to maximise the information-carrying capacity of the watermark. Thus each data bit in the watermark induces a pattern of modulation in a block or window of the spectrogram. The pattern of modulations is applied in one sense to encode a "one" bit and in the opposite sense to encode a "zero" bit. Bits are encoded at regular time intervals, namely at a regular horizontal spacing T in the spectrogram.

It is possible that there will be short segments of the covertext in which it is impossible to hide a watermark such as the silent sections of an audio file.

It is therefore essential that each data bit affects as long a section of stegotext as possible. In the present embodiment two approaches to this problem are used.

Figure 4 of the accompanying drawings shows in diagrammatic form one of these approaches. In this approach the spectrogram modulation patterns for adjacent bits overlap. In Figure 4 each rectangle K represents a copy of the modulation pattern. Each spectrogram modulation pattern K is x time units wide and y frequency units high, y being the full height of the spectrogram. In the present embodiment x is 32 and T = 5. Thus when the first 32 column wide block of the power spectrogram of the covertext is modulated by a key of the same size and the key is then stepped by T (5 columns) then the initial five columns of the covertext will remain only modulated by the corresponding five columns of the key. In the next iteration of the modulation the columns 6 to 37 of the covertext will be modulated by the key so that columns 6 to 32 will have been modulated twice. At the third iteration columns 6 to 10 of the first block are left with only the double modulation, but columns 11 to 32 are modulated for a third time, while columns 33 to 37

will receive their second modulations and columns 38 to 42 their first modulation. This sequence is repeated for the entire length of the covertext. The values of x and T can of course vary over a wide range. For example x can be 256, and T can be 10.

The second approach is to apply an error-correction code to the message bits to spread the effect of each bit still further in time.

A convolution encoder shown in Figure 5 of the drawings is used to spread the effect of each input bit over a longer section of the music in a way which reduces the memory requirements in the decoder as compared to the use of a longer key. The data stream to be encoded is input on a line (30) to a shift register which in this embodiment consists of three D-type flip flops (31, 32 and 33). The clock (clk/2) is provided on a line (34). An output switch (35) which is flipped at a clock rate (clk) is connected to the outputs of a pair of exclusive-OR gates (36, 37) so as to select between one of two exclusive-OR combinations of the bits in the shift register formed by the three flip flops. In the present embodiment the upper exclusive-OR gate (36) is connected

to all three bits of the shift register and the lower gate (37) to bits 0 and 2. This encoder is specified by the two-dimensional matrix [111; 101], where the first row of the matrix corresponds to the shift register connections made to the upper exclusive-OR gate, and the second row of the matrix corresponds to the connections made to the lower exclusive-OR gate (37). The patterns of connections can be expressed in polynomial form, with coefficients from the set {0, 1}. In this case the polynomials are $X^2 + X + 1$ (gate 36) and $X^2 + 1$ (gate 37).

In this encoder each input bit affects six consecutive output bits (the total number of entries in the matrix) and the output bit rate is twice the input bit rate (the number of rows in the matrix). Such a code is called a "rate ½ code". The entries in each row of the matrix, the "generator polynomials", have to be chosen carefully. Only codes whose rate is the reciprocal of an integer can be used in the present embodiment. It will be appreciated that this restriction is only caused by the type of encoder used in the present embodiment and has no other relevance. Thus if another

form of error correction coding was to be used the restriction need not apply. The identity code, which passes to its output unchanged, is specified by the matrix [1].

The code is called "convolutional" because it can be implemented using a convolution function as follows. The input data bits are first interspersed with zeros according to the code rate. For example, suppose the original data are (1011). The data are interspersed with zeros to obtain (1000101) so that the data is now at half the original rate. Convolving these data with (111011) which is the above encoder matrix written as a single row, yields (111022212111). Taken modulo-2 this is (111000010111). The modulo-2 operation performs the function of the exclusive-OR gates (36 and 37) in the encoder. Thus a four-bit sequence has been encoded into a twelve-bit code word. In general an n-bit sequence is encoded into a (2n + 4) -bit code word.

Whilst a convolutional code has been described it will be appreciated that many other suitable types of error-correcting code can be used. Such codes include Reed-Solomon code, BCH codes, Golay codes, Fire codes,

Turbo codes, Gallagher codes and Mackay-Neal codes.

Synchronisation encoding is carried out before convolution encoding. Thus, synchronisation flags are inserted into the encoded stream of data bits. Thus synchronisation of the encoded bit stream is achieved by inserting a start flag into it. The flag pattern is a "0" followed by five "1"s. To ensure that this pattern does not otherwise occur in the data stream, any sequence of four "1"s has an extra zero bit inserted after it. This is known as "zero stuffing". The stuffed zeros are removed by the decoder. This procedure gives a penalty of six bits per start flag plus an overall reduction in data rate by just over three percent. Those skilled in the art will realise that many alternative methods for this are possible.

Returning now to Figure 6 of the drawings this shows how the two processes just described are incorporated in the encoder. Thus again (10) indicates the covertext to be encoded and K indicates a current window being processed and K-1 indicates the previous window. At the multiplier (41) the extracted window K is multiplied with an analysis window function (42) so that the extracted

window tapers at each of its ends. In circuit (43) the fast Fourier transform of the extracted window modified by the analysis window function is obtained. Rectangular polar conversion is carried out in circuit (44) so as to generate the required power spectrogram. This power spectrogram is modified as already described in circuit (45), which corresponds to circuit 12 of Figure 2, with the phase component of the spectrogram remaining unchanged.

To complete the generation of the stegotext, polar rectangular conversion is carried out in circuit (46); inverse fast Fourier transform is taken in circuit (47) and the synthesis window function (48) is used to multiply the output of the inverse fast Fourier transform circuit (47) at (49). Finally the overlapped windows are added at (50) to generate the stegotext indicated at (15).

It will be appreciated that it is desirable to have a number of different watermarks available. Basically the possibility of using a number of different watermarks makes it considerably harder for an attacker to decode, remove or falsify a hidden message without knowledge of

which mark is to be used.

In the present embodiment the word "key" is used to refer to a particular member of a family of watermarks. Again in the present embodiment the keys are generated pseudo-randomly, and any one key is determined by a single integer which is used as a seed. This is the seed input shown in Figure 1 of the drawings.

In the present embodiment the key is an array K(t,f) of spectrogram modulation values where t and f are integer indices and $-1 \le K(t,f) \le +1$. K(t,f) is defined to be zero outside the range $-X/2 \le t < X/2$ and $0 \le f < Y$. Let the spectrogram of the covertext be G(t,f) and the spectrogram of the stegotext be H(t,f). Let d_i represent the data bits to be encoded, where $d_i = \pm 1$ (rather than 0 or 1). In the interests of simplicity error-correction coding is ignored. Then the encoding algorithm is given by

$$H(t,f) = G(t,f)\prod_i e^{sd_iK(t-iT,f)} \qquad \ldots(1)$$

and hence, given suitable choices of branch cut,

$$\log H(t,f) = \log G(t,f) + s\sum_i d_iK(t-iT,f) \qquad \ldots(2)$$

where s is a real constant which determines the encoding strength. In equations (1) and (2) G and H are complex but K is real. Thus, by equation (1), argH=argG. The watermark is therefore encoded in the power spectrum, and the phases of the original spectral components are preserved.

It will be appreciated that the design of the key is of paramount importance in the generation of a stegotext which is robust against attack. Thus design consideration for the key will now be described in detail.

A key that consists simply of a white noise pattern, where each cell in the key is independently and identically distributed, is attractive for many reasons. It is computationally easy to generate and has the maximum possible information-carrying capacity. In general it has low correlation with the covertext and has a single narrow autocorrelation peak. Experiments have shown that it can be robust to a wide variety of manipulations of the audio file while still being encoded with sufficiently low strength to be inaudible. However by manipulating the stegotext using a group delay attack

on the spectrogram in which individual rows are shifted left or right at random, it is with the spectrogram resolution already given possible to arrange the group delay parameters so as to shift the rows by more than one column. This destroys any correlation between stegotext and key. It appears to be impossible to select a spectrogram resolution that simultaneously gives perceptually satisfactory construction of the stegotext and robustness against all forms of this group delay attack.

Moreover the stegotext can be resampled so that all frequencies increase by say 5% (less than one semitone), and the text shortens in time by the same factor. The effect of this on the spectrogram is to stretch it vertically and shrink it horizontally. This procedure is shown diagrammatically in Figure 7 where 15A represents the original stegotext and 15B is the altered stegotext. It can be seen that very few of the cells will still coincide; along the frequency axis cells with $f \geq 20$ will not overlap at all with their previous positions. The correlation function is again destroyed.

The first of these two problems, namely stretch in

one dimension, can be overcome by modifying the key so that it contains repeating columns. Experiment shows that repeating each spectrogram column twelve times is sufficient to ensure that the group delay required to destroy the correlation function has a perceptually unacceptable effect on the stegotext. The cost is in reduced information-carrying capacity: the autocorrelation peak of the key is wider and lower, and so a greater encoding strength is required for a given robustness.

The second problem can be overcome by exhaustive search. The correlation function can be evaluated at a range of different resampling rates and, by finding the one that gives the strongest correlation, determine by what factor the file has been resampled. Unfortunately, it is possible to resample a stegotext in such a way that the pitches change but the overall time remains constant, or so that the pitches stay constant but the overall time changes. This latter process is common in broadcast applications where, for example, it is desired to make a piece of music fit exactly a given slot. There is therefore a two-dimensional space of possibilities to

search: the stegotext may have been arbitrarily stretched in frequency and/or time. If the key has been modified to include repeating columns as above, the autocorrelation function is wide and hence the range of possible time stretches need only be sparsely sampled; nevertheless, the computational burden is great.

However the present invention provides a solution to this problem. Considering carefully the effect on the key of a stretch relative to some fixed origin, it can be seen that the relative effect of the stretch is constant over the key; it is the absolute effect that varies and which gives rise to the problem above. In the present embodiment the key pattern is modified so that higher spatial frequencies are filtered out further from the origin.

For the purposes of the following discussion the covertext, stegotext and key will be considered as images in the log spectrogram domain. When referring to "frequencies" this will mean spatial frequencies in these images, not frequencies in the underlying audio.

First consider the problem in one dimension. Let $f(t)$ be a sine wave, $f(t)=\sin \omega t$. Squashing by a factor

α, gives $g(t)=\sin \alpha\omega t$. The phase angle ϕ between these sine waves is given by $\phi=\alpha\omega t-\omega t=\omega t(\alpha-1)$. Insisting that the correlation between $f(t)$ and $g(t)$, calculated in a suitably-chosen interval around t, exceed some threshold is equivalent to bounding the phase angle ϕ so that $|\phi|<\phi_0$. Thus there is a constraint on ω in terms of $t:|\omega|<\phi_0/(\alpha-1)t$, or, where α is chosen to be the greatest stretch to which resistance is required, $1/\omega>C|t|$ for some positive constant C. In view of this relationship, it is simpler to talk in terms of the timescale τ of a sinewave, where $\tau=1/\omega$.

It is now possible to specify the frequency content of a function that correlates well with itself when stretched. It must contain no frequency components with timescales shorter than a threshold timescale $\tau=C|t|$, where the constant C sets the degree of stretch resistance desired. Such a function can be obtained by suitable filtering of a white noise signal. A low-pass filter is required whose cutoff frequency varies inversely with t. Such a filter will hereinafter be referred to as a "swept" filter.

As already described in the present embodiment keys

corresponding to consecutive data bits are overlapped. In order to minimise overlap between the frequency components present at a particular point in time due to one copy of a key and those due to previous or subsequent copies, a high-pass filter is also applied to the key. The overall effect is therefore that of a band-pass filter. The cutoff frequency of the high-pass filter is swept so as to match the low-pass characteristics of the adjacent keys. This is demonstrated in the graph of Figure 8. The "bandwidth" Δ, which is constant in terms of timescale, is given by $\Delta=CT$, T being the interval between consecutive applications of the key to the covertext. Figure 9 shows how the copies of the key for four consecutive bits d_0, d_1, d_2 and d_3 overlap.

An example of the result of applying such a swept band-pass filter of the type just described to a white noise signal is shown in figure 10.

Similarly, a two-dimensional key can be generated from a two-dimensional white noise pattern. A filter with characteristics varying as described above is applied separately in each dimension. After filtering, the data values are passed through a non-linear function to

enforce the condition $-1 \leq K(t,f) \leq +1$; in this embodiment, sine is used.

An example of the resulting pattern appears in the power spectrogram of Figure 11. Here the axes are 'time' and 'frequency' (audio frequency is now meant): these match the axes of the spectrogram of Figure 3 to which the key will be applied. The origin in the time direction is in the middle of the key, whereas that in the frequency direction is at the top. The right hand column in Figure 11 has a similar function to that of the scale column in Figure 3.

The present embodiment applies band-pass filtering not only along the x axis but also along the y axis, although since copies of the key are not overlapped in that direction just low-pass filtering would suffice. It is possible to increase the information-carrying capacity of the key by using a low-pass rather than a band-pass filter.

If the value of the constant C in the above equation is increased, the key generated becomes resistant to greater stretches. The high-pass and low-pass filter characteristics become closer together and so the pass

band of the filter becomes narrower. This reduces the information-carrying capacity of the key. There is thus a tradeoff to be made between stretch resistance and information-carrying capacity. In the present implementation C=0.15 (pixels per cycle) per pixel and the key so generated works satisfactorily under stretches of up to about ±6% in either the time or the frequency directions. It will be seen that the above definition of C refers to pixels. In the present context the term pixel has different meanings when considering filtering in the horizontal direction of the spectrogram and filtering in the vertical direction.

In the horizontal direction the term pixel is being used to mean the time interval between columns of the spectrogram. When considering filtering in the vertical direction the term pixel is used to mean the difference in frequency of two adjacent rows of the spectrogram.

Thus in Figure 11 a horizontal pixel is approximately 23 millisecs. whereas in the vertical direction it is approximately 22Hz.

Thus in using the formula $\tau > C|t|$ for the passband of the low pass filter τ is measured in pixels per cycle and

t represents the x or y coordinate of the point of the spectrogram under consideration measured in pixels as hereinbefore defined from a reference point or origin. In Figure 11 the reference point is at the centre of the top edge of the picture. This reference point is chosen so as to correspond to zero frequency. It is possible to select other reference points but the zero frequency condition is preferable.

In the present embodiment the key is generated in response as described to a seed integer in a manner which is itself known. Thus in key generator 1 a Tausworthe generator of uniformly distributed random numbers is used and the required 1-dimensionally Gaussianly (i.e. normally) distributed numbers are generated by the Box-Cox method. Both these procedures are fully described in "Principles of Random Variate Generation" by John Dagpunar, in the series Oxford Science Publications, published by Clarendon in 1988. The swept filtering is carried out computationally.

The peak correlation of a typical key with itself after a range of stretches in both frequency and time is shown in Figure 12. The values have been normalised so

that the peak autocorrelation of the key is 1.

When a two-dimensional correlation between the key and the stegotext is calculated, it is found that the correlation peak can move slightly away from the line y=0 when the stegotext has been stretched in frequency. For this reason the present embodiment uses a two-dimensional correlation; the values of the function at small offsets in the y direction are added together to form a one-dimensional function to pass to the bit synchroniser.

Having described the basic steps and principles of encoding a covertext in accordance with the present invention, Figure 13 shows a block diagram of an encoder.

As in previous figures (10) represents a covertext, in the present embodiment music, which is to be encoded and (15) represents the final stegotext.

In circuit (51) the covertext is transformed into a log abs spectrogram.

The spectrogram so generated is supplied to a FFT circuit (51) where the received spectrogram is clocked by a clock (52) into a spectrogram buffer. The FFT circuit (51) carries out the overlapping segmentation of the input spectrogram and the windowing function described in

Figure 6. The clock (52) ensures that the content of spectrogram buffer (53) represents in spectrogram form a quantity of music equal to the length of the key, in the present embodiment 256 or 32 columns.

The data to be encoded is supplied at (55) to a circuit (54) for adding, as already described, synchronisation flags and for carrying out zero stuffing.

The output of circuit (54) is supplied to the convolution encoder (56) which corresponds to the encoder described with respect to Figure 5 and which has the requisite polynomials supplied to it at (57).

The key matrix is supplied to the encoder at (58) to a circuit (59) where the key matrix is converted into a set of values which can be directly multiplied into the spectrogram held in the spectrogram buffer (53). These values are in the form of two matrices, one for encoding a zero bit and the other for encoding a one bit. These matrices are the antilog of the key and the reciprocal of the antilog of the key. The operation of multiplying these matrices into the held spectrogram is equivalent to adding or subtracting in the log spectrogram domain.

The degree of strength with which the key modulates

the contents of the buffer (53) is determined by an input (60). This input corresponds to the real constant s in equation 2.

The two matrices are selectively multiplexed with the contents of the spectrogram buffer (53) as indicated at (61), the selection being made in accordance with the output of the convolutional encoder (56) so that the music stored in buffer (53) is encoded with a single bit of data. The contents of the buffer (53) are shifted along by one clock period for each bit written into the IFFT circuit (62) so that the main encoding loop is executed once for each bit written.

The output of the IFFT circuit (62) is applied to an anti-clip buffer (63). This is to ensure that the data read from circuit (62) is not clipped when the data is written out as a music file. If clipping is imminent an amplitude modulation curve is generated to reduce the volume of the output gradually so that clipping is just averted. The volume is increased to normal, again gradually, when it is safe to do so.

Finally the output from anticlip circuit (63) is output as the stegotext (15).

Figure 13 also includes a scrambler 65. Many possible scramblers can be used but a typical one is described in the V32 standard of CCITT. The inclusion of a scrambler is optional, as is the convolutional encoder.

The previous description has for reasons of simplicity described the use of a single key. It will of course be appreciated that more than one key can be used, each key having been generated by a different seed integer. Additionally multiples of the key or keys can be used to watermark a stegotext. In the above described embodiment the multiple is "1" so that wherever a multiple of a key is mentioned it is implicit that the multiple can be "1" i.e. the key remains unchanged apart from its sign.

The actual way in which two or more different keys are used to watermark a stegotext or to retrieve the watermark code are entirely analogous to the embodiments described in this specification. Thus if there is more than one key which key is multiplied into the spectrogram at (61) at any one time will be determined in accordance with the data to be encoded. If multiples other than ± 1 are used to modulate the spectrogram more than 1 bit can

be encoded each time. Decoding will, of course, utilise the same set of multiples.

Having described an embodiment of an encoder according to the present invention attention will now be turned to the problem of decoding a stegotext which may have undergone compression or stretching to recover the coded data.

The previously described key is capable of dealing with distortions of the stegotext which involve stretches of the stegotext in either the vertical or horizontal directions of ±6%.

In order to deal with cases where the stegotext has undergone greater stretch than the ±6% allowed for by the key a number of approaches are possible.

The present embodiment of a decoder utilises an approach which involves direct correlation and this will be now described in detail.

Having discussed the characteristics of the key which is used in the encoder of Figure 13 to modulate the power spectrum of the covertext it will be appreciated that when decoding a stegotext to extract the watermark data the data bits can be identified by correlating the

key with the log power spectrogram of the stegotext. If the stegotext has not undergone an attack or has not otherwise being stretched or compressed there will be a clear correlation between the stegotext and the key at those log elements which have been modified in accordance with the data.

The main underlying principle of the decoder to be described is that of calculating the correlation function between the key and the stegotext. The correlation is carried out in the log power spectrogram domain so that the magnitude of each of spectrogram element, measured in dB is calculated. During encoding the key was multiplied or divided into the spectrogram of the covertext. This procedure in the log domain corresponds to respective addition or subtraction operations. The correlation of the log power spectrogram of the stegotext with a key will therefore extract the data as positive and negative peaks in the correlation waveform, the respective peaks corresponding to "1" and "0" bits.

As the method of carrying out the correlation operation uses fast Fourier transforms the spectrogram can be considered as a two-dimensional image with the

correlation process matching two images namely the stegotext and the key. The two-dimensional correlation is carried out in such a way as to allow the watermark to be found at arbitrary time shifts as well as small offsets in frequency. This is important in music as the correlation peak in the frequency direction can shift from zero when the music has undergone frequency stretching.

Turning now to the actual embodiment of the decoder shown in Figure 14, the stegotext (15) to be decoded is input to FFT circuit (71). This circuit is analogous to the circuits (41), (42) and (43) of Figure 6 and circuit (51) of Figure 13. Thus in circuit (71) the stegotext is point-wise multiplied by a windowing function comparable to the windowing function disclosed at (51,52) of Figure 6. The windowed stegotext is Fast Fourier Transformed and then converted into the log power spectrogram domain.

The output of circuit (71) is supplied to a spectrogram buffer (73). As with the spectrogram buffer (53) of Figure 13, buffer (73) holds a length of music in spectrogram form. The content of the buffer (73) are

sectioned into blocks corresponding in size to the original blocks which were modulated to generate the stegotext.

If it were known that the stegotext had not been tampered with by stretching or compression then given knowledge of the key and the original encoding clock rate it would be a relatively simple matter to correlate the key with the sample of the stegotext held in buffer (73). However as the stegotext may have undergone resampling or pitch-invariant time stretching these possibilities have to be compensated for. Thus it is important to obtain an estimate of the actual date clock rate and this is the function of the circuit (72). The nominal clock rate is known as it is provided as a parameter to the decoder having been used in the encoder. Thus the clock estimate circuit (72) receives not only the nominal clock rate but the output of the clock extraction circuit (82). The key stretch circuit (74) prepares 3 versions of the key. One is stretched so as to match music-up samples by 6⅔% and the other to match music-down samples by the same factor while the third is the unmodified key. Stretching the key involves sampling it along the time axis by a factor,

for example λ, and along the frequency axis by a factor $1/\lambda$. The two new versions of the key along with the original are used in the correlation block (75) to take part in a set of trial correlations using an initial sample of the stegotext in the spectrogram buffer (73).

In the present embodiment a block of music three times larger than the key is taken from the beginning of the stegotext and converted by circuit (71) in the manner already described into the two-dimensional power spectrogram form. The two-dimensional correlation function with each key is calculated in turn.

This two-dimensional correlation is carried out in circuit (75). As described the method of carrying out correlation utilises FFTs. Let the signals to be correlated by x and y and let their Fourier transforms be denoted by X and Y. Then the Fourier transform of the correlation function of x and y is $X\,Y^*$, where * denotes complex conjugate. Calculating $X\,Y^*$ and taking the inverse Fourier transform thus produces the desired result. Accordingly three trial decodings are carried out on the short initial segment using the correlation

function obtained with the key and with its two stretched versions. In each initial correlation an appropriately scaled version of the nominal data clock rate is taken as an initial estimate. The RMS value of the correlation function at the bit positions is evaluated and a refined data clock estimate corresponding to the largest RMS correlation value is then taken as the true data clock rate.

Circuit (75) shows that the Fourier transform of a block or segment of the spectrogram buffer is Fourier transformed in both dimensions at (76) and is multiplied by the conjugate of 2-d FFT of the key which was made by circuit (77).

The result of this multiplication is inverse transformed at circuit (79) to yield the two-dimensional correlation function of the sample of the stegotext with the key. It will be appreciated that the conjugate of the Fourier transform of the key need only be calculated once outside the main decoder loop once the clock rate has been established.

In circuit (79) the lower frequency components of this correlation function are summed to form a one-

dimensional correlation function. This one-dimensional correlation function is then overlapped with, and added to the end of the correlation function from the previous block. It will be appreciated that this procedure is necessary because of the way in which the key was applied to overlapping segments of the original covertext.

The output of (79) is supplied for a correlation function buffer (80). This buffer accordingly holds a series of values representing the peaks generated by the correlation between the key in the stegotext and the key as supplied by circuit (74). The content of the buffer are illustrated in Figure 15. The values in the buffer (80) are represented by the dark curve (150). The solid vertical lines are the times at which the bits have been encoded as determined by the clock extraction circuit (82).

These values are read from buffer (80) by a data slicing circuit (81) operating in response to the extracted clock as generated by a clock extraction circuit (82).

The output of the data slicing circuit (81) generates a series of values on either side of zero with

the positive values potential "+1"s and the negative values potential "-1"s.

The final stages of the decoder of Figure 14 are conventional.

The convolutional encoder shown in Figure 5 generates two output bits for each input code bit.

Therefore for every two bits present in the output of circuit (81) and stored at (83) a decision has to be made as to which bit is part of the desired code.

This function is carried out by a convolutional decoder (84) which, in its simplest form, looks at each potential output bit, and for each such bit considers all the possible values of the surrounding bits within a fixed window. This procedure is carried out in phase search circuit (85). The size of the window is a compromise between performance and amount of calculation. For example for a window encompassing ten values in the buffer a total of 1024 sequences have to be evaluated.

For each of the 1024 sequences the encoded value is calculated and the probabilities that values in the buffer over that window are calculated by adding or subtracting the relevant values in the buffer in

accordance with the relevant bit being +1 or -1.

The probabilities of all the 512 sequences which have a +1 in the position under consideration are added, and the other 512 sequences which have a zero in the relevant position are added. This gives a probability that the bit under consideration is a 1 or a 0.

This procedure is illustrated in Figure 16. In this figure (250) in a schematic representation of the values sliced by circuit (81). Win_i represents a ten-value window and V_i represents a pixel of interest. Win_{i+2} represents the next window in this sequence and V_{i+2} the next value. Finally V_i represents the result of the evaluation just carried out for the value of interest V_i and V_{i+2} the outcome of the next evaluation.

As shown in Figure 16, because the output of the convolutional encoder of Figure 5 gave 2 output bits for each code bit, the window is stepped at two bit intervals along the contents of the buffer. This procedure has to be carried out twice over the respective odd and even numbered values in the buffer. Two sequences are thus generated each having a probability associated with it and finally a selection is made on the basis of the

sequence which has the higher probability.

What has just been described is the simplest form of encoder/decoder.

However it may be advantageous to have some other ratio other than 2 to 1.

If for example the ratio was 4 to 1 it would be necessary to use four sequences in which a window was successively stepped evenly for values in the buffer, and to select the most probable output bit from these four sequences.

Equally there are other ways in which the code can be decoded which will be apparent to people skilled in the art such as a Viterbi decoder.

Decoder polynomials corresponding to those used at the encoding stage are supplied to the maximum-likelihood decoder 81 at 86, and finally added synchronisation bits and zeros added during zero stuffing are removed at (87) to leave the decoded data at (88).

The descrambler 89 is required only if the optional scrambler of the encoder is present.

It will be appreciated that in the foregoing specification the various embodiments of encoders and

decoders have been defined in terms of circuit elements such as "filter", "multiplier", "buffer", and "circuit" and so on. However apart from the actual recording or reproduction of a signal all these circuit elements can be replaced by appropriate software manipulation. Thus in particular the encoder described with respect to Figure 14 can be replaced in all its functional aspects by a general purpose computer receiving appropriate code. An example of such a code is given with regard to the generator of the matrix B_0 used in the decoder of Figure 21. Thus all the steps and blocks shown in Figures 14, 16 and 20 can have their functions carried out as software steps.

In the case of the decoder embodiments if there are to be used in individual systems which as well as decoding the stegotext produce the stegotext as an output, for example as music, then the decoders may well be in the form of integrated microprocessor(s) possibly employing very large scale integrated circuits.

CLAIMS

1. An encoder for encoding a covertext signal to generate a stegotext, the encoder comprising:

first transformation means for carrying out a Fast Fourier Transform and rectangular polar conversion of the covertext signal so as to transform the covertext signal into a log power spectrogram;

means for providing at least one key, the or each key being in the form of a two-dimensional pattern of predetermined size;

a multiplier for adding or subtracting in the log power spectrogram domain multiples of the key or multiples of one or more of the keys if there is a plurality of keys, to blocks of the transformed covertext signal;

means for controlling the addition or subtraction of the key or keys by the multiplier in accordance with data representing a desired code; and

second transformation means for carrying out polar rectangular conversion and inverse Fast Fourier Transformation of the modulated covertext signal to generate a stegotext.

2. An encoder according to claim 1, wherein there is a single key and the multiple of the key is one.

3. An encoder according to claim 1 or claim 2, wherein said first transformation means are adapted to segment the covertext into overlapping segments prior to carrying out the Fast Fourier Transformation and rectangular polar conversion.

4. An encoder according to claim 3, and including a multiplier for multiplying each segment of the covertext with a function so that each segment tapers at its ends.

5. An encoder according to claim 4, wherein the function is the square root of a raised cosine function.

6. An encoder according to any ones of claim 3 to claim 5, wherein said first transformation means are adapted to convert each segment into the log power spectrogram domain to generate blocks which are of the same length and have the same number of columns as the key.

7. An encoder according to claim 6, wherein each block of the log power spectrogram of the covertext signal is x columns wide, and the multiplier is adapted to apply the key to the blocks in steps of T, where T is an integer number of columns of the key, so that the key is applied at least in part, approximately 2 x/T times to

each spectrogram block.

8. An encoder according to claim 7, wherein each block is a power spectrogram which is 32 columns wide and 1024 bits high and the key is also 32 columns wide and 1024 bits high.

9. An encoder according to claim 8, wherein T equals five columns of a spectrogram block.

10. An encoder according to any one of the preceding claims, wherein said second transformation means are adapted to carry out polar rectangular conversion and inverse Fast Fourier Transformation on each modulated block of the covertext signal and to synthesise the resulting segments to generate the stegotext.

11. An encoder according to claim 9, wherein said second transformation means are adapted to multiply each segment resulting from the polar rectangular conversion and inverse Fast Fourier Transformation with a function so that each segment tapers at each end prior to synthesis to generate the stegotext.

12. An encoder according to any preceding claim wherein

each use of the key represents a single data bit.

13. An encoder according to any one of the preceding claims and including means for setting the amount of modulation produced by the key on a block of the transformed covertext to be approximately + or - 1dB.

14. An encoder according to any one of the preceding claims and including an error correction encoder for error correction encoding the watermark code data before the data is used to control the watermarking of the covertext.

15. An encoder according to claim 14, wherein the error correction encoder is a convolution encoder.

16. A method of encoding a covertext signal to generate a stegotext, the method comprising:

carrying out a Fast Fourier Transform and rectangular polar conversion of the covertext signal so as to transform the covertext signal into the power spectrogram domain;

providing at least one key, the or each key being in the form of a 2-dimensional pattern of predetermined size;

adding or subtracting in the log power spectrogram domain multiples of the key or multiples of one or more of the keys if there is a plurality of keys, to segments of the transformed covertext signal;

controlling the addition or subtraction of multiples of the key or keys at the addition/multiplication step in accordance with data representing a desired code; and

carrying out polar rectangular conversion and inverse Fast Fourier Transform of the modulated covertext signal to generate a stegotext.

17. A method according to claim 16, wherein there is a single key and the multiple of the key is one.

18. A method according to claim 16 or claim 17, comprising segmenting the covertext into overlapping segments prior to carrying out the Fast Fourier Transformation and rectangular polar conversion.

19. A method according to claim 14, and including multiplying each segment of the covertext with a function so that each segment tapers at its ends.

20. A method according to claim 19, wherein the function is the square root of a raised cosine function.

21. A method according to any ones of claims 14 to 20, wherein each segment of the covertext is converted into the log power domain to generate blocks which are of the same height and have the same number of columns as the key.

22. A method according to claim 21, wherein each block of the log power spectrogram of the covertext signal is x columns wide, and the multiplier is adapted to apply the key to the blocks in steps of T, where T is an integer number of columns of the key, so that the key is applied, at least in part, 2 x/T times to each spectrogram block.

23. A method according to claim 22, wherein each block is a power spectrogram which is 32 columns wide and 1024 bits high and the key is a 2-dimensional pattern of the same height and width.

24. A method according to claim 23, wherein T equals five columns of a spectrogram block.

25. A method according to any one of claims 16 to 24, wherein the transformation of the modulated covertext signal into the stegotext is by carrying out polar

rectangular conversion and inverse Fast Fourier Transformation on each modulated block of the covertext signal and synthesising the resulting segments.

26. A method according to claim 25, wherein each segment resulting from the polar rectangular conversion and inverse Fast Fourier Transformation is multiplied with a function so that each segment tapers at each end prior to synthesis to generate the stegotext.

27. A method according to any one of claims 14 to 26 and including setting the amount of modulation produced by the key to be approximately 1dB.

28. A method according to any one of claims 14 to 27 and including error correction encoding the data of the code prior to this data being used to control the addition or subtraction of the key into the power spectrum of the transformed covertext signal.

29. A method according to claim 28 wherein the error correction encoding uses a convolutional encoder.

30. A method according to any one of the claims 14 to 29 and including scrambling the code data used to modulate

the covertext signal.

31. A method according to any one of claims 14 to 30 wherein the covertext is a signal representing music.

32. A storage medium storing processor implementable instructions for controlling a processor to carry out the method of any one of claims 14 to 31.

33. An electrical signal carrying processor implementable instructions for controlling a processor to carry out the method of any one of claims 14 to 31.

34. A storage medium storing in readable format a stegotext encoded by the method of any one of claims 14 to 31.

35. A signal carrying a stegotext encoded by the method of any one of claims 14 to 31.

36. A decoder for decoding a stegotext signal watermarked by the method of any one of claims 14 to 31 to obtain the watermarking code, the decoder comprising:

transformation means for carrying out Fast Fourier Transformation and rectangular polar conversion of the

covertext signal so as to transform the stegotext signal into the log power spectrogram domain;

means for providing the key or keys with which the original covertext signal was encoded;

means for carrying out 2-dimensional correlation between the key or keys and blocks of the transformed stegotext signal which are of the same length as the key so as to generate a correlation function representing the correlation between the key and the stegotext; and

means for extracting the data bits representing the code from the correlation function.

37. A decoder according to claim 36 including means for estimating the clock period of the stegotext by generating at least two additional versions of the key, one version being stretched in time with regard to the nominal clock rate and the other being compressed in time;

means for carrying out sample correlations utilising the alternate versions of the key and the nominal clock rate;

means for selecting the most appropriate clock rate from the sample correlations;

means for carrying out correlation of the transformed stegotext with the key in accordance with the

clock as determined by the sample correlations to generate a correlation function of the key with the power spectrogram of the covertext and means for extracting from the correlation function so generated the data used to watermark the stegotext.

38. A method decoding a stegotext signal watermarked by the method of any one of claims 13 to 27 to obtain the watermarking code, the method comprising:

carrying out Fast Fourier Transformation and rectangular polar conversion of the stegotext signal so as to transform the stegotext signal into the log power spectrogram domain;

providing the key or keys with which the original covertext signal was encoded;

utilising the key or keys to obtain a 2-dimensional correlation function between the key or keys and sequential blocks of the transformed stegotext signal which are of the same length as the key or keys so as to generate a correlation function representing correlation between the key and the stegotext; and

means for extracting the data bits representing the code from the correlation function.

39. A method according to claim 38, including estimating

the clock period of the stegotext by generating at least two additional versions of the key, one version being stretched in time with regard to the nominal clock rate and the other being compressed in time;

carrying out sample correlations utilising the alternative versions of the key and the nominal clock rate;

selecting the most appropriate clock rate from the sample correlations;

carrying out correlation of the transformed stegotext with the key in accordance with the clock as determined by the sample correlations to generate the required correlation function; and

extracting from the correlation function so generated the data used to watermark the stegotext.

Application No: GB 0018491.1 **Examiner:** Martyn Dixon
Claims searched: all **Date of search:** 30 January 2001

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.S): H4R (RPTS,RPX); H4P (PDCFX,PDX,PEUM); H4F (FBB); G5R (RHB)

Int Cl (Ed.7): H04H (1/00); G11B (20/00); H04N (1/32,5/913,7/08)

Other: Online: EPODOC,WPI,JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
A,E	GB 2348028 A	(Tandberg)	
A	GB 2343818 A	(IBM)	
A	WO 96/42151 A	(The Dice Co)	
A	EP 0905967 A	(Digital Copyright Technologies) see especially figs 4-7 and col 23, lines 11-19	
A	EP 0891071 A	(Matsushita) see especially col 12, line 53	
A	EP 0828372 A	(NEC) see especially col 4, line 46	

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.